

TITLE

Marius Jonsson (Institutt for Vanskelig Fysikk, Oscars gate 19, 0352 OSLO, Norway)

<http://github.com/kingoslo/attaboy>

November 15, 2016

ABSTRACT

This is a report submission for the fourth project of «Computational physics» at the Institute of Physics, University of Oslo, autumn 2016.

INTRODUCTION

- > I Motivate the reader, the first part of the introduction gives always a motivation and tries to give the overarching ideas
- > I What I have done
- > I The structure of the report, how it is organized etc

METHODS

We used the Metropolis algorithm in conjunction with the Ising model to approximate the Boltzman distribution $P(E)$. The Metropolis algorithm works by constructing a Markov chain such that the stationary distribution is P (Robert, Christian; Casella, George (2004). Monte Carlo Statistical Methods). We know that a Markov chain is defined by its transition distribution. We know that detailed balance(hva) and ergodicity(hva) are sufficient but (but not required conditions) for existence and uniqueness of stationary distribution respectively.

Let's design a markov chain such that its stationary distribution is P . We will call this markov chain the Metropolis algorithm. We use the detailed balance criterion using P as the asymptotic distribution. This ensures that the resulting markov chain will have P as an asymptotic distribution.

$$p(i, j)P(j) = p(j, i)P(i) \quad \text{only if} \quad \frac{P(i)}{P(j)} = \frac{p(i, j)}{p(j, i)}$$

Now pick any conditional distribution $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ and a function $A : \mathbb{R}^2 \rightarrow \mathbb{R}$ such that

$$p(i, j) = g(i, j)A(i, j) \quad \text{only if} \quad \frac{A(i, j)}{A(j, i)} = \frac{P(i) g(j, i)}{P(j) g(i, j)} \quad (1)$$

Clearly, by fixing the transformation rules of A and g one uniquely defines p since p is the pointwise product of A with g . Let g be any function on \mathbb{R}^2 with non-compact support, and define A by

$$A(i, j) = \min \left(1, \frac{P(i) g(j, i)}{P(j) g(i, j)} \right), \quad (2)$$

which solves equation (1) (check yourself!). This is the selection rule for the Metropolis algorithm as you will find implemented in this project.

And sinec this markov chain is ergodic (Hjorth jensen), P is the only stationary distribution of of p (Robert, Christian; Casella, George (2004). Monte Carlo Statistical Methods). Let us consider how this is implemented in the case of the Ising model, where $A(i, j)/A(j, i)$ is taken to be the Boltzman factor.

In this project we assume we are working with a sequence of $L \times L$ matrices M_n where the elements $m_{u,v}$ are either 1 or -1 . Let J be some real number, we assign a quantity called the energy denoted by E which is computed by

$$E_n = -J \sum_{(i,j) \in S_L^2} (m_{ij}m_{i(j+1 \bmod L)} + m_{ij}m_{(i+1 \bmod L)j}), \quad (3)$$

letting $S_L^2 = \{(u, v) \in \{1, \dots, L\} \times \{1, \dots, L\}\}$, this ensures the sum in (3) is over all products of neighbouring elements of the matrix M including periodic boundary conditions. We take the target distribution P for the Metropolis algorithm to be the Boltzmann distribution. Hence $P(i) = (1/Z)e^{-\beta E_i}$, for $\beta = (k_B T)^{-1}$ and k_B the Boltzmann constant. We let $g(i, j)$ be the unknown function induced by by the selection rule A and Boltzmann factor when we flip one spin at a time. Bla bla

The algorithm was implemented in c++ by:

```
for (int k = 0; k < N; k++) { // sample iterations
    for (int i = 0; i < m; i++) { // loop matrix cols
        for (int j = 0; j < n; j++) { // matrix rows
            int u = rand_int(gen);
            int v = rand_int(gen);
            double dE = 2*A[u][v]*(A[u][mod(v+1,n)] +
                                   A[mod(u+1,m)][v] +
                                   A[u][mod(v-1,n)] +
                                   A[mod(u-1,m)][v]);
            if (exp(-beta*dE) > rand_double(gen)) {
                // selection criterion
                A[u][v] = - A[u][v];
                E += dE;
                M += 2*A[u][v];
            }
        }
    }
}
// sample if we believe we're at equilibrium
if (k > samplepoint) {
    avg[0] += E;
    avg[1] += E*E;
    avg[2] += abs(M);
    avg[3] += M*M;
}
```

We computed the exact result for $L = 2$ and compared them to the numerical results. The errors were reasonable, indicating that the algorithm was implemented correctly (see table 5). The expressions for the exact results will be presented in the results.

We would like to know when we have reached the asymptotic distribution P . Since we want to compute observables, this is useful because this ensures that we know we are sampling to our expectation values from the Boltzmann distribution at any given confidence level. To test this, we will use a conventional χ^2 statistical goodness-of-fit-test for the null hypothesis $H_0 : P(i)$ is the asymptotic distribution against $H_a : P(i)$ is not the asymptotic distribution. The first markov iteration for which this is the case should be taken

as the sampling point, because then we know that we are sampling from the Boltzmann distribution. We test this by at each step n by assembling the asymptotic distribution by bootstrapping (Berk Devoire) and drawing samples from the empirical distribution at step n . We gather data by simulating N times using the Metropolis algorithm. We let the algorithm continue for a long time, typically 10^6 or 10^7 Montecarlo cycles. We do this to ensure that the asymptotic distribution is the Boltzmann distribution. We store the measured energies at each monte carlo cycle n for all the N simulations. This means we have typically 10^6 or 10^7 candidates to the asymptotic distribution (the boltzmanm distribution). Since we did exactly N simulations, we have N observation from the candidates to the asymptotic distribution at each n . Each of these candidates are compared to asymptotic distribution. There exist statistical tests which can be used to determine when the sample is drawn from the asymptotic distribution for each n at any confidence level α . The most common is the chi-square goodness-of-fit test, which works as follows: At each n , divide cut the domain of the empirical distribution into k intervals I_j such that there are at least 5 observations in I_j for all j . Count the number of observation in each interval, and label it by n_j . Then compute the probability of an observation to fall into I_j from the empirical distribution. Call this probability p_j . It can be shown using the center limit theorem of dependent stochastic that the following observable has a known distribution:

$$X^2 = \sum_{j=1}^k \frac{(n_j - np_j)^2}{np_j} \sim \chi_v^2 \quad \text{under } H_0.$$

The distribution is the chi square distribution with ν degrees of freedom, where $\nu = k - 1$ under the null hypothesis that the empirical distribution is equal to the empirical. In the case that the null-hypothesis is false, the value of X^2 will be large, and when the value of X^2 is larger than the $\chi_{\nu, 1-\alpha}^2$ -quantile in the χ^2 -distribution, we know that the distributions are difference at confidence level α . We chose a confidence level of 0.5, which implies we will only accept distributions which are so similar to the asymptotic distribution, that the probability of evaluating a distribution violating the null more hypothesis more strongly is exactly 50%.

RESULTS AND DISCUSSION

We start by presenting our main result. By making simulating for a long time on the available cluster, Smaug, we were able to simulate the same markov chain exactly 248 times for each temperature $T = 2, 2.1, \dots, 2.4$ and for dimensions $L = 20, 60, 100, 140$ for exactly 10^6 each iterations each run. With 248 samples from the distribution of E for each iteration, we were able to construct the empirical distribution of E at each iteration, giving us in total $1.8 \cdot 10^7$ probability distributions of E which could be candidates to the Boltzmann distribution. We automated the Goodness of fit test described in the methods, to discriminate which of these distributions were equal to the Boltzmann distribution at confidence level 95%. Using this information, we found the optimal sampling point given starting conditions (see figure 1). Table 1 gives the ideal starting points for these conditions. This could be further generalized using the method of maximum likelihood.

We found that for starting matrix of uniformly distributed spins, the temperature was unimportant for the ideal sampling point for $L = 20$, but was increasing for each temperature. An increase in temperature of 1 would reduce the sampling point by approximately one fourth. However, sampling point was strongly positively correlated with temperature for larger values of L . For example, for $L = 140$ the ideal sampling point was shifted by exactly four order of magnitudes by a shift in temperature of 0.5 (see table 1). This is reflected in the mean ideal sampling point given by $S = 1.3 \cdot 10^4$ with standard error of $3.1 \cdot 10^4$. We found that choosing an ordered starting matrix reduced the ideal sampling point by order of magnitudes. Now, at $L = 140$ and $T = 2$, the ideal sampling point was at $2 \cdot 10^1$, a reduction of four orders of magnitude. Naturally, since the lowest energy states are states where all spins are ordered, there was a linear increase in sampling point as a function of temperature for our temperature range (see table 2).

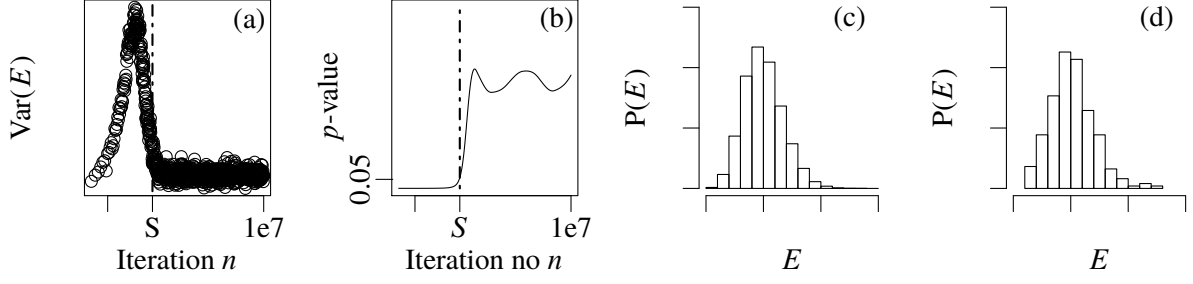


Figure 1: Subfigure (a) displays variance of some empirical distribution as a function of iteration number n for some temperature and starting matrix. We found that at a confidence level of 95%, the Markov chain arrived at the steady state distribution at the point marked by the dashed line. This is at $n = S$. (b) The p -values of the χ^2 -test notices that steady state has been reached and shifts from low to high at confidence level at the suggested sampling point S . (c) The steady asymptotic distribution at $n = 10^7$. This is the Boltzmann distribution. (d) The empirical distribution at confidence level 95% marked by the dashed line sampled at the point $n = S$. Notice that the empirical distribution is similar to the asymptotic distribution at this point since we are 95% certain that they are equal.

L	$k_B T$				
	2	2.1	2.2	2.3	2.4
20	$2.7 \cdot 10^2$	$2.4 \cdot 10^2$	$1.9 \cdot 10^2$	$1.1 \cdot 10^2$	$1.3 \cdot 10^1$
60	$5.3 \cdot 10^3$	$3.6 \cdot 10^3$	$2.5 \cdot 10^3$	$7.6 \cdot 10^2$	$9.0 \cdot 10^1$
100	$1.7 \cdot 10^4$	$1.7 \cdot 10^4$	$6.5 \cdot 10^3$	$9.2 \cdot 10^2$	$8.3 \cdot 10^1$
140	$1.3 \cdot 10^5$	$5.8 \cdot 10^4$	$1.5 \cdot 10^4$	$1.3 \cdot 10^3$	$1.2 \cdot 10^2$

Table 1: Points where $p > \alpha$, choses sampling points. Random starting matrix

L	$k_B T$				
	2	2.1	2.2	2.3	2.4
20	$7.0 \cdot 10^0$	$1.2 \cdot 10^1$	$2.5 \cdot 10^1$	$6.1 \cdot 10^1$	$6.1 \cdot 10^1$
60	$1.5 \cdot 10^1$	$2.1 \cdot 10^1$	$4.9 \cdot 10^1$	$7.1 \cdot 10^2$	$7.1 \cdot 10^2$
100	$2.0 \cdot 10^1$	$3.3 \cdot 10^1$	$9.3 \cdot 10^1$	$2.0 \cdot 10^3$	$2.0 \cdot 10^3$
140	$2.0 \cdot 10^1$	$3.8 \cdot 10^1$	$1.1 \cdot 10^2$	$3.4 \cdot 10^3$	$3.4 \cdot 10^3$

Table 2: Points where $p > \alpha$, choses sampling points. Static starting matrix

But these results would complement if we could produce an upper bound on the estimates $\langle E \rangle$, $\langle M \rangle$, C_V and χ as a function of data samples. Since we flip L^2 spins each iteration, the correlation between iterations will be weak, and we will assume it to be zero to simplify calculations. Since we are sampling after the iteration point $n = S$ the sampling variables $\text{Var}(E_i) = \sigma^2$ and $\langle E_i \rangle = \mu$ for some real numbers μ and σ^2 for all i . We start by computing the variance and mean of $\langle E \rangle$:

$$\text{Var}(\langle E \rangle) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n E_i\right) = \frac{1}{n^2} n \sigma^2 = \frac{\sigma^2}{n}, \quad \langle E \rangle = \frac{1}{n} \sum_{i=1}^n \langle E_i \rangle = \frac{1}{n} \sum_{i=1}^n \langle E_i \rangle = \mu.$$

We can see from figure 2, that for temperatures near $T = 2$, E_i is approximately normal, and thus a standardized squared observation is chi-square distributed with one degree of freedom and hence with variance equal to 2 (\dagger). We will also need Cauchy's formula (*) together with the assumed independence

between observations (**) to deduce that:

$$\begin{aligned}
\text{Var}(C_V) &= \frac{1}{k_B^2 T^4} \text{Var} \left(\langle E^2 \rangle - \langle E \rangle^2 \right) = \frac{1}{k_B^2 T^4} \text{Var} \left(\frac{1}{n} \sum_{i=1}^n E_i^2 - \left(\frac{1}{n} \sum_{i=1}^n E_i \right)^2 \right) \\
&\stackrel{(*)}{=} \frac{1}{2k_B^2 T^4 n^4} \text{Var} \left(\sum_{i=1}^n \sum_{j=1}^n (E_i - E_j)^2 \right) \stackrel{(**)}{=} \frac{(2\sigma^2)^2}{2k_B^2 T^4 n^4} \sum_{i=1}^n \sum_{j=1}^n \text{Var} \left(\left(\frac{E_i - E_j}{2^{1/2}\sigma} \right)^2 \right) \\
&\stackrel{(\dagger)}{=} \frac{(2\sigma^2)^2}{k_B^2 T^4 n^2}.
\end{aligned}$$

It was straightforward to compute the expectation of C_V . Under the same assumptions as previously we found that

$$\begin{aligned}
\langle C_V \rangle &= \frac{1}{k_B T^2} \langle \langle E^2 \rangle - \langle E \rangle^2 \rangle = \frac{1}{k_B T^2} \left\langle \frac{1}{n} \sum_{i=1}^n E_i^2 - \left(\frac{1}{n} \sum_{i=1}^n E_i \right)^2 \right\rangle \\
&\stackrel{(*)}{=} \frac{1}{2k_B T^2 n^2} \left\langle \sum_{i=1}^n \sum_{j=1}^n (E_i - E_j)^2 \right\rangle = \frac{2\sigma^2}{2k_B T^2 n^2} \sum_{i=1}^n \sum_{j=1}^n \left\langle \left(\frac{E_i - E_j}{2^{1/2}\sigma} \right)^2 \right\rangle \\
&\stackrel{(\dagger)}{=} \frac{\sigma^2}{k_B T^2}.
\end{aligned}$$

Unfortunately, we did not compute the same quantities for $\langle M \rangle$ and χ , however by experiment, we found these to be same order of magnitude as $\langle E \rangle$ and C_V . Using the above results one can proceed to compute various quantities. One quantity of interest could be the standard error. For example, we found that in order to obtain a standard error smaller than 5%, it would be necessary to simulate approximately 10^6 iterations beyond the ideal sampling point. Our exact findings are summarized in table 3.

L	$k_B T$				
	2	2.1	2.2	2.3	2.4
20	$1.1 \cdot 10^4$	$1.5 \cdot 10^4$	$2.1 \cdot 10^4$	$2.6 \cdot 10^4$	$2.2 \cdot 10^4$
60	$1.1 \cdot 10^5$	$1.4 \cdot 10^5$	$2.0 \cdot 10^5$	$3.0 \cdot 10^5$	$1.7 \cdot 10^5$
100	$2.9 \cdot 10^5$	$3.8 \cdot 10^5$	$5.6 \cdot 10^5$	$7.6 \cdot 10^5$	$4.4 \cdot 10^5$
140	$5.8 \cdot 10^5$	$7.2 \cdot 10^5$	$1.1 \cdot 10^6$	$1.1 \cdot 10^6$	$8.3 \cdot 10^5$

Table 3: Required iterations before 5 percent standard error

Despite our first findings, it was mandatory to simulate for $L = 20$ and two temperatures $T = 1, 2.4$. We plotted the means of energy $\langle E \rangle$ and magnetization $\langle M \rangle$ and number of accepts per monte carlo cycles. We found that in the case of $T = 1$, the ordered matrix started its markov chain near the asymptotic distribution. The means $\langle E \rangle$ and $\langle M \rangle_n$ was nearly constant. This was also the case for the acceptance numbers. In the case of a random starting point it appeared that approximately 10^3 monte carlo cycles was necessary to achieve constant acceptance rate. The estimators $\langle E \rangle$ and $\langle M \rangle$ had an error smaller than 1% after roughly 10^4 iterations. In either case, the asymptotic distribution $P(E)$ appear Boltzmann with parameters $(\beta, n) = (1, 400)$. In this case, the variance of energy was 9.48 yielding a heat capacity per spin of 0.024 (in units of k_B). In the case of $T = 2.4$, the acceptance rates of both an ordered and random starting point were approximately constant with standard deviation roughly 60. The distribution appeared roughly Boltzmann with parameters approximately $(\beta, n) \approx (100, 400)$. This yielded a variance of 3226.65 and consequently a heat capacity per spin of 1.40. The errors on the values of $\langle E \rangle$ and $\langle M \rangle$ was smaller than 1% at approximately $n = 10^5$. See figure 2 for the plots.

We found exact expressions for the partition function and Boltzmann distribution for our model in the case $L = 2$. Using the definition, we obtained the partition function:

$$Z(\beta) = 12 + 2(e^{-8J\beta} + e^{8J\beta})$$

and thus the n th momentums of energy E and magnetization M are given by

$$\langle E^n \rangle(\beta) = \frac{2}{Z} (8^n e^{-8J\beta} + (-8)^n e^{8J\beta}), \quad \langle M^n \rangle(\beta) = \frac{1}{Z} (4^n e^{8J\beta} + 4(2)^n + 4(-2)^n + (-4)^n e^{8J\beta})$$

respectively. From these we derived the expressions for expected energy and expected magnitude of magnetization

$$\langle E \rangle(\beta) = \frac{2^4}{Z} (e^{-8J\beta} - e^{8J\beta}) \quad \text{and} \quad \langle |M| \rangle(J\beta) = \frac{2^3}{Z} (2 + e^{8J\beta}).$$

Then again, use these, it was straight forward to compute the heat capacity at constant volume and magnetic susceptibility since they were

$$C_V = \frac{1}{kT^2} \sigma_E^2 = \frac{2^7}{kT^2 Z} \left[(e^{-8J\beta} + e^{8J\beta}) - \frac{2}{Z} (e^{-8J\beta} - e^{8J\beta})^2 \right]$$

$$\chi = \frac{1}{kT} (\langle M^2 \rangle - \langle |M| \rangle^2) = \frac{2^5}{kTZ} \left[(1 + e^{8J\beta}) - \frac{2}{Z} (2 + e^{8J\beta})^2 \right]$$

In the case $T = 1$, we found that these estimate resonated with the numerical results, given in table 5.

Fysikk bla bla

Observable	Exact:	Numerical, $n = 10^8$	Error
$\langle E \rangle$	-7.9839283437467596	-7.98379632	$-1.3 \cdot 10^{-4}$
$\langle M \rangle$	3.9946429309943987	3.9945928	$5.0 \cdot 10^{-5}$
C_V	0.12832932745714487	0.129378400754	$-1.0 \cdot 10^{-3}$
χ	0.01604295806490974	0.0162110021882	$-1.7 \cdot 10^{-4}$

Table 4: BLa bla

Previously, we noted that C_V could potentially have a pole near the point T_0 . It turns out that theoretically, we would expect a phase transition near the point T_0 . We will call this point T_C , the critical temperature. We were able to produce point estimates of T_C . According to the project description, the critical temperature T_C and our $T_C(L)$ for a finite grid of spins is related by

$$T_C(L) - T_C = aL^{-1/\nu}.$$

We let $\nu = 1$ and suggest point estimators for T_C and a by:

$$\hat{T}_C = \frac{1}{n} \sum_{i=1}^4 (T_C(L_i) - \hat{a}L_i^{-1}) \quad \text{with} \quad \hat{a} = \frac{1}{12} \sum_{i=1}^4 \sum_{j \neq i}^4 \frac{T_C(L_i) - T_C(L_j)}{L_i^{-1} - L_j^{-1}}$$

By entering the simulated quantities $T_C(L_k)$ we found that

By simulating for 15 hours on 248 cores we found estimates of $\langle E \rangle, \langle |M| \rangle, C_V$ and χ for $L = 20, 60, 100, 140$ for 201 temperatures in the interval $[2, 2.4]$. We found that for $L = 20$, the expected energy as a function of thermal energy was approximately linear (see figure 3). However, as the periodicity increased to say $L = 140$, the thermal energy increased near the temperature $T_0 = 2.25$. We discovered a magnetic transient

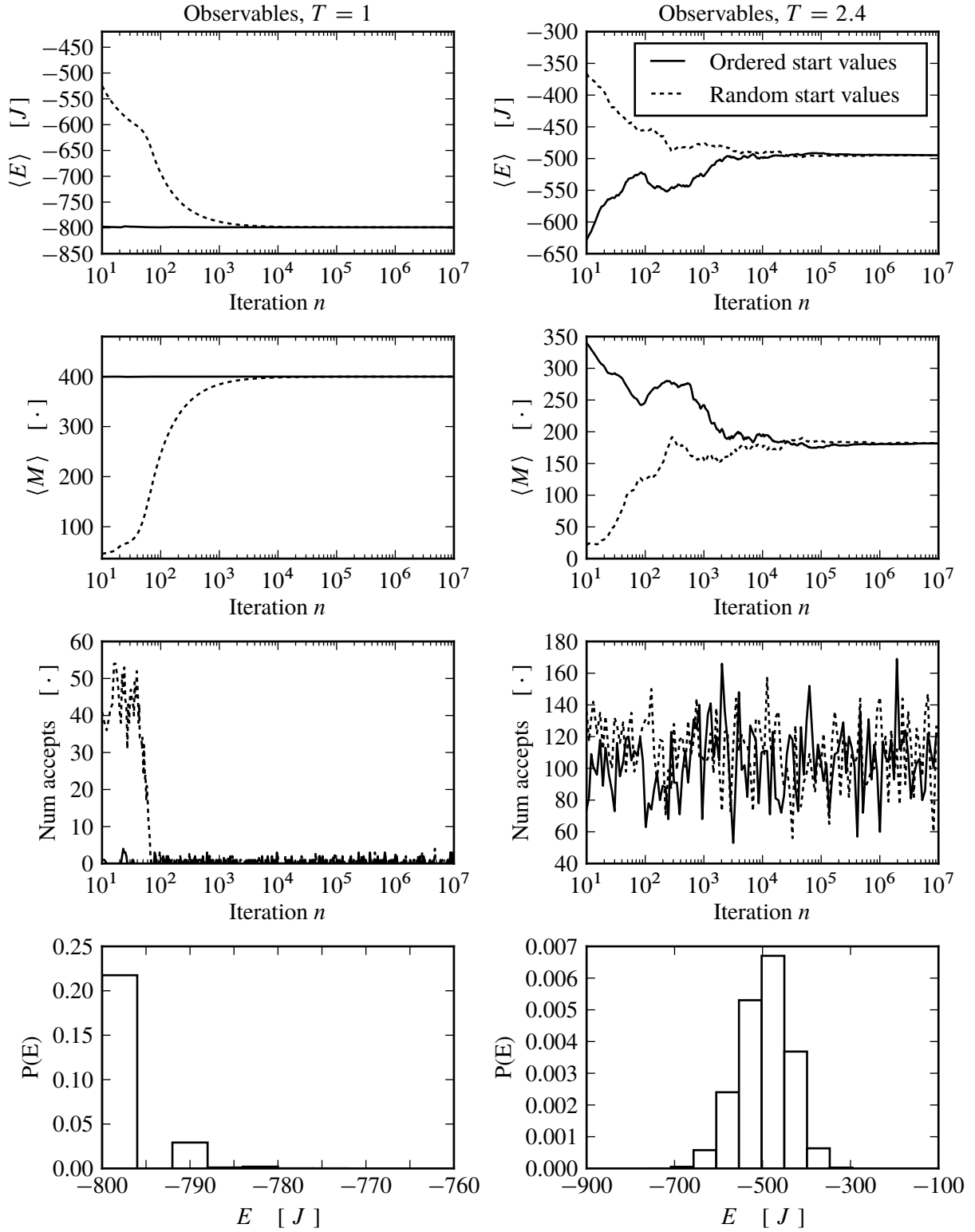


Figure 2: BLa bla Observables plotted against thermal energy for exactly three values of L . The solid line represent observables at $L = 140$, whilst the dotted, dashed and double-dot-dashed lines represent $L = 100, 60.20$ respectively. The solid line represent observables at $L = 140$, whilst the dotted, dashed and double-dot-dashed lines represent $L = 100, 60.20$ respectively. BLa bla

Exact:	Numerical, $n = 10^6$	Error
2.26918531421	2.27017857143	$-9.9 \cdot 10^{-4}$

Table 5: BLa bla

at the same temperature to lower levels of magnetization. By considering a periodicity of $L = 140$, the difference in magnetization was nearly an order of magnitude relative to $L = 20$ at T_0 . The magnetic susceptibility and heat capacity had global maxima near T_0 . It appeared C_V may have a pole near T_0 in the limit $L \rightarrow \infty$ (see figure 3).

> samplingspunktanalyse > regresjonsformel (bruk random som faktor) vent til resultat kommer > $L = 20$, $T=1, 2.4$. Plot mean E mean M , accepts both random and ordered. > Analytical expr 2x2. Periodic boundry
> Write code to compute numerical. $T = 1$ give table
fys > Plott E, M X C_V for $L=20,60,100,140$, t in 2 til 2.4 > Beregn kritisk temperatur > Beregn $P(E)$ for $T = 1, 2.4$ og $L=20$. Finn variansen fra fordelingene.

Diskusjon > samplingspunktanalyse > regresjon > $L = 20$, $T=1, 2.4$. Accepts both random and ordered: How do accepts behave as function of T ? (INTERESSANT!) Study how long before equilibrium. Can we use this to find equilibrium time? How many cycles before equilibrium? > Kan vi se tegn til faseovergang? > Diskuter oppførselen til $P(E)$ som funksjon av temp. for $T = 1, 2.4$, $L = 20$ > Numerics: 2x2 How many iterations n is necessary for good fit?

> I Present your results

> I Give a critical discussion of your work and place it in the correct context.

> I Relate your work to other calculations/studies

> I An eventual reader should be able to reproduce your calculations if she/he wants to do so. All input variables should be properly explained.

> I Make sure that figures and tables should contain enough information in their captions, axis labels etc so that an eventual reader can gain a first impression of your work by studying figures and tables only.

CONCLUSION AND PERSPECTIVES

> I State your main findings and interpretations

> I Try as far as possible to present perspectives for future work

> I Try to discuss the pros and cons of the methods and possible improvements

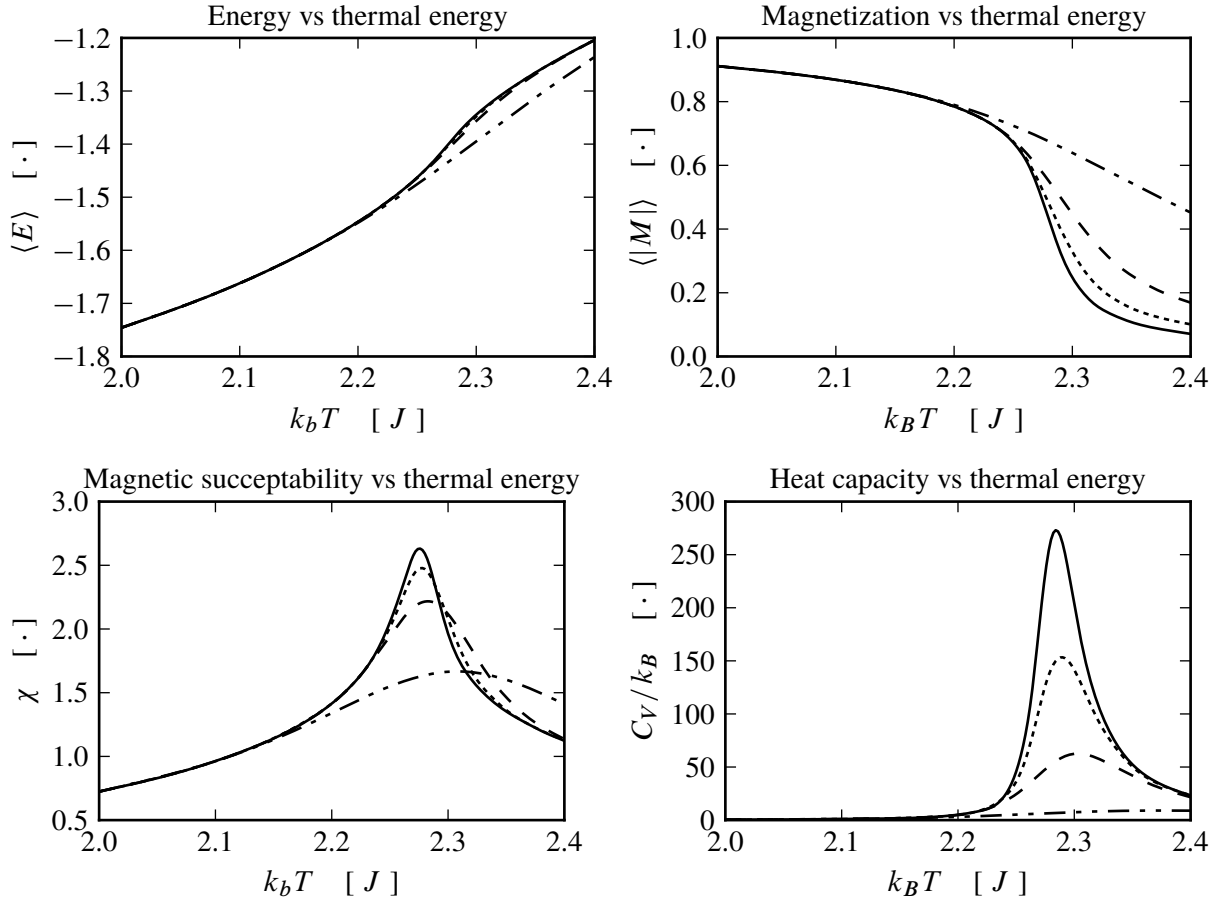


Figure 3: Observables plotted against thermal energy for exactly three values of L . The solid line represent observables at $L = 140$, whilst the dotted, dashed and double-dot-dashed lines represent $L = 100, 60, 20$ respectively.